From *Cognitive Consistency: A Fundamental Principle in Social Cognition.*
Edited by Bertram Gawronski and Fritz Strack.

9

Implicit Ambivalence

Richard E. Petty
Pablo Briñol
India Johnson

Various contemporary attitude theorists have made a distinction between explicit and implicit measures of attitudes. Explicit measures tap into the evaluations that people can self-report. In contrast, implicit measures attempt to assess attitudes that come to mind automatically, without requiring self-report. People are invariably aware of the attitudes reported on explicit measures, but they might or might not be aware of the attitudes assessed with implicit measures (Petty, Fazio, & Briñol, 2009a).

Although the implicit–explicit attitude distinction has become very popular recently (see reviews by Gawronski & Payne, 2010; Petty, Fazio, & Briñol, 2009b; Wittenbrink & Schwarz, 2007), it has been around in one form or another for a long time. For example, in their classic treatise on persuasion, Hovland, Janis, and Kelley (1953) defined *attitudes* as "implicit responses" that are "sometimes unconscious" and "oriented toward approaching or avoiding a given object" (p. 7). Attitudes were contrasted with *opinions*, which are "verbal answers that one covertly expresses to (oneself)" (p. 8). These private opinions were further distinguished from public opinions that could be susceptible to social desirability motives. Today, one might say that their use of the term *attitude* referred to the underlying (implicit) evaluative association, whereas *opinion* referred to the explicit or deliberative evaluation. Although, in Hovland's day, all that could be measured were explicit

evaluations, today it is possible to assess automatic evaluative associations as well (e.g., Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998). Hovland and colleagues (1953) assumed that explicit and implicit attitudes would generally correspond, but available research now makes it clear that implicit and explicit measures can reveal quite different (even opposite) evaluations.

Of most relevance to this volume is that the new implicit measures have led to new developments with respect to classic notions of cognitive consistency. For example, whereas early studies showed that people could experience imbalance (Heider, 1958) when their explicit cognitions about different objects were not in accord with each other (e.g., "People I like do not like the same things that I do"), more current research suggests that the same can be true when there is imbalance among evaluations assessed with implicit measures (e.g., Gawronski, Walther, & Blank, 2005). Importantly, *implicit imbalance* can occur even when everything is balanced at the explicit level and implicit balance can occur even if explicit cognitions are imbalanced (Greenwald et al., 2002). For these situations to arise, of course, implies that implicit and explicit evaluations differ from each other for at least one of the elements in the cognitive system.

When implicit and explicit evaluations of a *single* attitude object differ, we refer to this situation as implying a state of *implicit ambivalence* toward the attitude object. *Explicit ambivalence* has been much studied and refers to a situation in which people endorse both positive and negative aspects of a given attitude object (e.g., "I like ice-cream because of its great taste, but I don't like it when I think of the high calories"; Kaplan, 1972). This ambivalence is explicit in that people report both positive and negative features of the attitude object on self-report scales, and when asked also report being ambivalent with respect to the attitude object (Priester & Petty, 1996; Thompson, Zanna, & Griffin, 1996). In contrast to explicit ambivalence, people with implicit ambivalence do not label themselves as ambivalent, because at the explicit level they do not endorse conflicting evaluations of the object. Rather, implicit ambivalence occurs when people are either unaware of the evaluative conflict (e.g., aware of their positive but not their negative reactions to the object), or are aware of having both positive and negative reactions, but deny that one reaction is valid or believe it stems from some source other than their personal views (e.g., from media bias). Thus, as we explain in more detail shortly, people can be implicitly ambivalent toward an attitude object even if there is no explicit ambivalence. However, despite not labeling their reaction as ambivalent, we argue that implicit ambivalence is consequential.

We begin the chapter by very briefly describing some work on implicit imbalance—a type of inconsistency involving multiple attitude objects. Then, we turn to the main focus of this chapter, which is implicit ambivalence. We address issues such as how and why implicit ambivalence occurs, and what its consequences are. Unlike the principle of balance that seems to operate

at the primary level of thinking, implicit ambivalence requires a consideration of secondary, metacognitive labels, along with primary cognitions. In the last part of the chapter, we introduce recent research suggesting a means to reduce implicit ambivalence.

BALANCE

Explicit Balance

Balance principles (Heider, 1958; see also Abelson et al., 1968; Festinger, 1957) apply to many aspects of attitudes and persuasion. Heider proposed that three elements in a cognitive system (e.g., self, other, object) could either be in an evaluatively balanced state (e.g., a person agreeing about the value of some object with another person who is liked, or disagreeing about the value of some object with another person who is disliked) or in an imbalanced state (e.g., a person disagreeing about the value of some object with another person who is liked, or agreeing about the value of some object with another person who is disliked). Balanced triads can involve any three cognitive elements that are associated in some way (e.g., "I like myself and I have nice things" is balanced, whereas "I dislike myself and I have nice things" is not).[1] Importantly, Heider held that imbalanced systems were unpleasant and unstable, and tended to move toward balance. This theory can account for why a person would come to like a political candidate more after he or she is endorsed by a favored celebrity. The theory holds that imbalance can lead to attitude change toward the candidate (or the endorser) in the direction of producing balance.

Implicit Balance

Recently, Greenwald et al. (2002) proposed a unified theory of implicit social cognition in which the essential ingredients of Heider's theory of psychological balance were applied to automatically activated cognition (see Cvencek, Greenwald, & Meltzoff, Chapter 8, this volume). The core elements in the unified theory are self, group, and valence (attitude), but the framework would presumably be applicable to all of the elements addressed by Heider's original theory. In the unified theory, the self can be automatically associated with one or more (nonvalenced) group concepts (e.g., "I am a Democrat"), and the self and group elements can be automatically associated with a particular valenced node (e.g., positive–negative). Greenwald and his colleagues examined the coherence among these three constructs using implicit measures. For example, in one study with female students examining gender identity, the authors measured three types of associations: self–gender (gender identity), self–valence (self-attitude), and gender–valence (gender attitude). In line with the predictions of balance theory, they found that ingroup attitudes were the result of a multiplicative function of

the strengths of ingroup identity and self-attitude (Greenwald et al., 2002); that is, the stronger the automatic associations between the self and female and the self and good, the stronger the automatic linkage between female and good. Interestingly, the balance effects at the implicit level were stronger than any balance effects at the explicit level.

Other studies have taken the implicit balance notion beyond self–group links. For example, research by Brunel, Tietje, and Greenwald (2004) revealed that Macintosh computer users showed stronger automatic evaluative associations for that brand relative to PC users, as revealed both in positivity (brand–good) and in self associations (brand–self). Indeed, objects, products, and brands can be part of the self-concept and contribute to defining one's identity (e.g., Aaker, 1999; Belk, 1988; James, 1890; Kleine, Kleine, & Kernan, 1993; Wicklund & Gollwitzer, 1982). Finally, some relevant research has examined the possibility of a more dynamic approach to balance by manipulating the relationship between two concepts and assessing their consequences for the relationship with the third concept (Gawronski, Bodenhausen, & Becker, 2007; Walther & Trasselli, 2003).

In research particularly relevant to persuasion, Horcajo, Briñol, and Petty (2010) showed that asking participants to think about persuasive messages can also lead to associated changes on automatic measures through a process of spreading activation guided by the attainment of psychological balance. In one study, participants were asked to generate arguments in favor of or against including more vegetables in their diet, linking vegetables to either goodness or badness. Then, they completed an implicit measure (the Implicit Association Test [IAT]) designed to assess the automatic link between vegetables and the self, as well as a measure of implicit self-esteem (Greenwald et al., 2002). Consistent with the idea of balance, participants showed more automatic self–vegetable associations after thinking about why vegetables were good rather than bad, and these effects were only apparent for those with high implicit self-esteem (i.e., those with stronger automatic associations between the self and good rather than bad). For individuals with low implicit self-esteem there was a nonsignificant tendency toward stronger associations between self and vegetables after thinking about the negative consequences of consuming vegetables. According to balance theory, self-esteem should matter, because balance occurs when good people have good objects linked to the self, but bad people have bad objects linked to the self.

In a second study, Horcajo and colleagues (2010) provided further evidence for the dynamic relationship between the three concepts under examination (self, vegetable, and attitude). In this study participants received false feedback about their self-concept to increase or decrease the perceived linkage between the self and vegetables. Specifically, after completing a relevant IAT they were told that their self-concept was strongly associated with either vegetables or animals. Then, the impact of this induction was assessed on an implicit measure of attitudes toward vegetables (i.e., the link between vegetables and their valence), as moderated by scores on implicit self-esteem (i.e.,

the link between the self and its valence). It was predicted and found that the false feedback increasing the self–vegetable linkage led to more favorable implicit attitudes toward vegetables, but only for those with relatively high implicit self-esteem. As in the prior study, self-esteem matters, because if vegetables become linked to the self, the situation is balanced when vegetables are seen as good only so long as one likes oneself. If one dislikes oneself and vegetables become linked to the self, balance occurs if vegetables are seen as bad.

The research by Horcajo et al. (2010) makes several advances over previous work on implicit balance. The most important of these extensions is to add to the research moving implicit balance from a relatively static to a more dynamic view (see also Gawronski et al., 2005; Langer, Walther, Gawronski, & Blank, 2009); that is, whereas the initial research on implicit balance by Greenwald and colleagues (2002) focused on the examination of implicitly measured constructs (e.g., the self, a group, and valence), the Horcajo et al. (2010) research took a more experimental approach to examine implicit balance in which one link was manipulated and the effects on the other links were observed. Overall, the evidence suggests that regardless of whether one takes a relatively static or a more dynamic approach, the results are highly consistent with the notion of implicit balance (see Cvencek et al., Chapter 8, this volume, for additional discussion).

AMBIVALENCE

Although imbalance results from inconsistency with respect to multiple attitude objects, and ambivalence results from inconsistency regarding just one object, there are number of similarities between these two kinds of inconsistency. First, both balance and ambivalence approaches assume that people can simultaneously hold incompatible beliefs, attitudes, feelings, and behavioral tendencies regarding the self, others, and objects. Second, literature in both paradigms has documented that such conflicts are typically aversive. Virtually every consistency theory holds that internal discrepancies are unpleasant and often result in negative affect (e.g., Abelson & Rosenberg, 1958; Higgins, 1987; Newcomb, 1978; Osgood & Tannenbaum, 1955). Third, both approaches postulate that people's attempts to deal with internal discrepancies in one way or another sometimes lead to attitude change in the service of restoring consistency. For balance, we have seen that change can occur at both explicit and implicit levels.

Along with these similarities, there are also notable differences. First, as noted earlier, balance effects operate at the primary level of cognition. *Primary thoughts* are those that involve our initial associations of some object with some attribute, such as "I am female" or "I am a good person." Following a primary thought, people can also generate other thoughts that occur at a second level, which involve reflections on the first-level thoughts (e.g., "I

am not sure how good a person I really am"). *Metacognition* refers to these second-order thoughts, or our thoughts about our thoughts or thought processes (Briñol & DeMarree, in press; Jost, Kruglanski, & Nelson, 1998; Petty, Briñol, Tormala, & Wegener, 2007). As we explain shortly, unlike implicit imbalance, implicit ambivalence involves secondary levels of cognition, as well as primary cognition.

Another distinction between balance and ambivalence is that the resolution of imbalance appears to occur spontaneously, whereas ambivalence (more like cognitive dissonance; see Festinger, 1957) requires people to search for solutions. Perhaps the most common approach to dealing with ambivalence is enhanced thinking or processing of information relevant to the object of discrepancy (e.g., Abelson et al., 1968; Aronson, 1969; Festinger, 1957; Hass, Katz, Rizzo, Bailey, & Moore, 1992; Maio, Bell, & Esses, 1996). By considering additional information, and especially attitude-consistent material (Clark, Wegener, & Fabrigar, 2008), individuals presumably hope to gain enough information to resolve or minimize their ambivalence, or at least the subjective discomfort that results from it (e.g., Hänze, 2001; Hodson, Maio, & Esses, 2001; Jonas, Diehl, & Bromer, 1997; Katz, Wackenhut, & Hass, 1986).[2]

The Metacognitive Model

As noted earlier, explicit ambivalence occurs when people have mixed evaluative reactions to any one attitude object; that is, they recognize that there are both good and bad features of the object. Although there are a number of ways to conceptualize ambivalence (e.g., see Priester & Petty, 1996; van Harreveld, van der Pligt, & de Liver, 2009), we use a model of attitude structure called the metacognitive model (MCM; Petty, 2006; Petty & Briñol, 2006; 2009; Petty, Briñol, & DeMarree, 2007). The MCM makes a number of assumptions about attitudes. First, in accord with a view of attitudes as stored representations (e.g., Fazio, 1995, 2007; Fiske & Pavelchak, 1986) rather than momentary constructions (e.g., Schwarz, 2007; Wilson & Hodges, 1992), the MCM holds that attitude objects can be linked in memory to global evaluative associations, and these associations can vary in their accessibility (see Fazio, 1995, for a review). For many attitude objects, one evaluation (relatively positive or negative) is dominant and represents the integration of knowledge about the object. For example, the top panel of Figure 9.1 shows a person with an explicitly positive attitude toward smoking. In such situations, this evaluation would come to mind upon encountering the attitude object, though the speed at which this occurs can vary (e.g., see Bargh, Chaiken, Raymond, & Hymes, 1996; Fazio, Sanbonmatsu, Powell, & Kardes, 1986).

Second, and more importantly, in concert with idea that the positivity and negativity underlying attitudes are separable (e.g., Cacioppo, Gardner, & Berntson, 1997), the MCM holds that attitude objects can sometimes be linked in memory to evaluative associations of opposite valence (see also de Liver, et al., 2007; Petty, Tormala, Briñol, & Jorns, 2006). In such cases, a per-

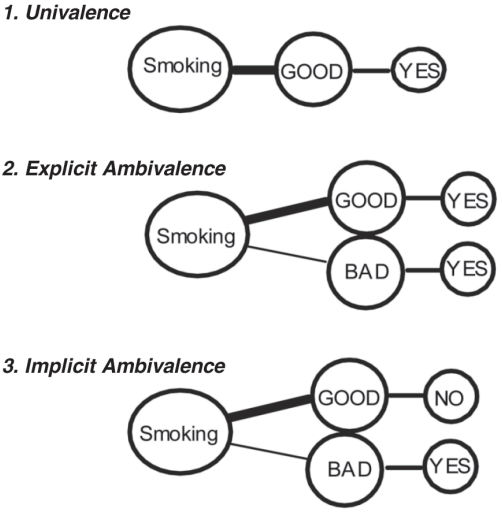

FIGURE 9.1. Metacognitive model of attitude structure depiction of univalence (Panel 1), explicit ambivalence (Panel 2) and implicit ambivalence (Panel 3). Adapted from Petty, Briñol, and DeMarree (2007). Copyright 2007 by Guilford Publications, Inc. Adapted by permission.

son would have explicit ambivalence regarding the attitude object, because the object is linked to both positivity and negativity. Panel 2 in Figure 9.1 depicts a person who views smoking as having both positive and negative aspects and thus is explicitly ambivalent. In this illustration, the link to good is stronger than the link to bad, perhaps because negativity toward smoking is more recent and less practiced. It is important to note, however, that not all attitude objects are expected to have a bivalent structure.

Third, the most unique feature of the MCM, and the one that gives the model its name, is the assumption that people can tag their evaluative associations as true or false, or held with varying degrees of confidence. In this way, the model builds on prior research on attitude certainty (e.g., Gross, Holtz, & Miller, 1995) and on metacognition more generally (Jost et al., 1998; see Petty, Briñol, Tormala, et al., 2007). These validating metacognitions can vary in the strength of their association to the linked evaluation, and the strength of the attitude–validity link will determine the speed and likelihood that the perceived validity of an evaluation will be retrieved along with the evaluation itself. Most notably, perhaps, the MCM goes beyond the idea that attitude validation is solely an online process and contends that perceived validities, like the evaluations themselves, can be stored for later retrieval (cf. Gawronski & Bodenhausen, 2006).

Finally, the MCM concurs with research on cognitive negation, which suggests that untagged evaluations are presumed to be true unless evidence

against them has been generated. This proposal is analogous to Gilbert's suggestion that information initially held as true needs to be tagged as "false" to be disbelieved (Gilbert, 1991; Gilbert, Krull, & Malone, 1990; Gilbert, Tafarodi, & Malone, 1993). Only if the false tag is retrieved will a person who disbelieves an assertion or has abandoned an attitude recognize it as false. Otherwise, the person will act upon the belief or attitude as if it were true. The accumulated research suggests that successful negation is quite difficult, though not impossible (Deutsch, Gawronski, & Strack, 2006). Negation is particularly difficult with respect to automatic evaluation when people first believe something and only at some subsequent point in time come to learn that it is wrong (Gregg, Seibt, & Banaji, 2006; Petty, Tormala, Briñol, & Jarvis, 2006). If a false tag is applied immediately to a new statement, it can be effective in correcting immediate evaluations on both explicit and implicit measures (Peters & Gawronski, 2011), though at some later point in time even immediate negations can be forgotten, leaving people to think that the negated assertion is valid (e.g., Priester, Wegener, Petty, & Fabrigar, 1999; Ross, Lepper, & Hubbard, 1975).

Feeling ambivalent is associated with a number of important consequences. As noted earlier, subjective ambivalence, like other forms of inconsistency, is often experienced as an unpleasant state (Newby-Clark, McGregor, & Zanna, 2002). In this regard, ambivalence differs from *indifference*. Although both ambivalent and indifferent individuals could express a "neutral" or moderate attitude on a bipolar scale (Kaplan, 1972), for an ambivalent person various positive and negative reactions are jointly activated, whereas for an indifferent person there are few positive and negative reactions. An indifferent person is better characterized as having a "non-attitude" that is not associated with as much discomfort as ambivalence (Converse, 1970). In addition to ambivalence being associated with discomfort, the more ambivalence one experiences regarding an object, the less functional one's attitude becomes in orienting one's behavior (Armitage & Conner, 2000; Sparks, Harris, & Lockwood, 2004). Consistent with this idea, people with ambivalent (vs. univalent or indifferent) attitudes tend to be slower to report their attitudes (Bargh, Chaiken, Govender, & Pratto, 1992) and are more sensitive to context effects (Tourangeau, Rasinski, Bradburn, & D'Andrade, 1989). Given that subjective ambivalence tends to be a negative state, people are motivated to address it. For example, the motivation to reduce ambivalence can lead people to pay careful attention to information that might help them undermine or resolve their ambivalence (e.g., Maio et al., 1996).

Implicit Ambivalence

Although classic work on ambivalence considers situations in which people hold both positive and negative evaluations that are considered valid (see Panel 2 in Figure 9.1), the MCM holds open the possibility that people can

have two opposite accessible evaluations, but one is seen as valid, whereas the other is rejected (see Panel 3 of Figure 9.1). A denied evaluation can be a past attitude (e.g., "I used to like smoking, but now it is disgusting") or an association that was never endorsed but is nonetheless salient due to one's culture (e.g., the media depicting smoking as glamorous; cf. Monteith, Devine, & Zuwerink, 1993). In such cases (Panel 3 of Figure 9.1), the MCM refers to the attitude structure as one of *implicit ambivalence* (Petty, Tormala, Briñol, & Jarvis, 2006). Even though people do not explicitly endorse opposite evaluations of the same attitude object, they can nevertheless feel uncomfortable about such attitude objects, without recognizing the specific source of the conflict (see Epstein, 2003; Petty & Briñol, 2009; Rydell, McConnell, & Mackie, 2008). The ambivalence is not explicit, because people do not endorse both positive and negative evaluations. Indeed, people can report being confidently unambivalent in such cases (e.g., "I am sure that I am not ambivalent about smoking"; Clarkson, Tormala, & Rucker, 2008). However, the underlying structure of the attitude could still suggest conflict, because the attitude object is associated with both automatically activated positivity and negativity.

According to the MCM, implicit ambivalence is suspected whenever there is a discrepancy in the valence of an attitude uncovered by a deliberative (explicit) versus automatic (implicit) measure. This is because in situations where people have opposite evaluative associations but reject one of them, a deliberative measure would be influenced primarily by the endorsed association, whereas a measure of automatic evaluation would tap into both endorsed and unendorsed associations. In the example depicted in Panel 3 of Figure 9.1, a deliberative attitude measure would reflect an unfavorable assessment of smoking, but an implicit measure would be more positive. The prediction of implicit ambivalence from implicit–explicit discrepancies stands in contrast to theories that assume that implicit and explicit measures tap into independent "dual attitudes" (see Wilson, Lindsey, & Schooler, 2000) that are stored in separate brain regions (see DeCoster, Banner, Smith, & Semin, 2006), stem from qualitatively different processes (see Rydell & McConnell, 2006; Rydell, McConnell, Mackie, & Strain, 2006), and operate in distinct situations (see Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997).

Implicit–Explicit Discrepancies as a Source of Implicit Ambivalence

If discrepancies between implicit and explicit measures tap an implicit ambivalence, then such discrepancies should be associated with some discomfort regarding the attitude object that would lead to efforts at resolution. As noted earlier, one well-known consequence of the discomfort that emerges from explicit ambivalence is that it leads to enhanced processing of relevant information in a presumed attempt to resolve the ambivalence.

In one study testing the notion that explicit–implicit discrepancies could likewise lead to enhanced information processing (Briñol, Petty, & Wheeler, 2006), undergraduates' self-evaluations were assessed with both automatic (IAT; Greenwald & Farnham, 2000) and deliberative (Rosenberg, 1965) self-esteem measures, then the absolute value of the difference between the two standardized measures was calculated as the index of discrepancy. Next, participants were exposed to either a strong or weak message about eating vegetables that was framed as self-relevant or not. An example of a strong argument in favor of vegetable consumption was that vegetables have more vitamins than most vitamin supplements on the market, making them particularly beneficial during exam and workout periods. The gist of one weak argument in favor of vegetables was that vegetables are becoming more popular for wedding celebrations because they are colorful and look beautiful on plates. The degree to which participants processed the message information was assessed by examining the extent to which the quality of the arguments affected postmessage attitudes toward vegetables. People's attitudes should be more affected by the quality of the message arguments when they are thinking carefully about information, than when they are not thinking carefully (see Petty & Cacioppo, 1986).

The results of this study revealed that when the message was framed as self-relevant, the extent of explicit–implicit discrepancy interacted with argument quality to affect attitudes. The greater the discrepancy, the more participants' attitudes differentiated strong from weak arguments. However, when the same strong and weak messages were framed as irrelevant to the self (i.e., the message was said to be about the properties of vegetables rather than the self), discrepancy did not interact with argument quality to predict attitudes. This suggests that explicit–implicit discrepancies do not lead to motivation to process all information—only information that is relevant to the object for which the discrepancy exists. Furthermore, the direction of the discrepancy (i.e., whether implicit or explicit self-esteem was greater did not further moderate the results.

Attitude Change as a Source of Implicit Ambivalence

In addition to examining discrepancies that already exist, we have also investigated implicit–explicit discrepancies created in the laboratory. Specifically, in a series of studies, we examined whether changing a person's attitude from one valence to another would produce implicit ambivalence. According to the MCM, explicit attitude change from one valence to another could produce implicit ambivalence, because one evaluative association (the old one) is now rejected and a new, opposite evaluative association is created, producing a pattern like that in Panel 3 of Figure 9.1. In this depiction, an explicit measure would primarily assess the new attitude, whereas an implicit measure would still show some influence of the old attitude, resulting in a discrepancy. In support of this notion, one study (see Gregg, Seibt,

& Banaji, 2006) found that when an initially disliked group subsequently became likable, explicit attitudes toward the group were influenced to reflect the new information, but implicit attitude measures still reflected the original dislike, resulting in an implicit–explicit discrepancy. According to the MCM, an implicit measure was especially likely to reflect the old attitude after explicit attitudes changed to the extent that the old attitude was initially highly accessible, such as when it was based on many rehearsals or extensive thought (see Wyer, 2010, for supporting evidence). In order to effectively reduce the impact of an initial attitude, it helps if the initial attitude is explicitly rejected at the same time that a new one is formed (Peters & Gawronski, 2011) and the information on which the new attitude is based is thought about carefully (Dempsey & Mitchell, 2010). These factors strengthen the new attitude relative to the old one.

To examine whether attitude change could lead people to show implicit ambivalence toward the attitude object that was the target of change, Petty and colleagues (2006) conducted a series of studies in which explicit attitudes were changed from one valence to another and information processing was then assessed. In one of the studies in this line of research, undergraduate students were first classically conditioned to like or dislike a target individual (Staats & Staats, 1958). Then, participants received information about the target individual's attitudes on several important topics. The attitudinal information was designed to get the participant either to like or dislike the target by having the target agree or disagree with the participant on several attitude issues (Byrne, 1971). In some conditions, this information had the same valence as the conditioning manipulation, so that no attitude change would occur, and in other conditions the information was opposite in valence to the conditioning. In the latter situation, an earlier study showed that individuals rejected their initial evaluation based on conditioning and adopted a new evaluation based on the similarity information. Next, participants were told that the target person was a candidate for a job at their university. To evaluate the candidate, participants were provided with either a strong, impressive résumé or a weak, unimpressive one. The key result was that attitudes toward the target as a job candidate were more influenced by résumé quality among participants whose attitudes were changed than among those whose attitudes toward the candidate had not been changed; that is, despite having the same current explicit attitude toward the target, people who held this attitude as a result of changing it engaged in greater information processing than those who had held this attitude all along. It was as if the changed group was attempting to resolve some underlying ambivalence regarding the candidate. This result held even though participants did not report being any more explicitly ambivalent (i.e., having any more mixed reactions) about the candidate as a function of changing their attitudes.

In a conceptual replication of our procedure, Rydell et al. (2008) examined whether people who changed their attitudes experienced more discomfort than people who did not change their attitudes, and whether this dis-

comfort was responsible for the enhanced information processing observed. In their research, participants were first exposed to either positive or negative associative information (based on a subliminal word prime) about a target person. Then, participants received positive or negative verbal information about the behaviors of the target. As in the Petty et al. (2006) design, in some conditions the verbal information had the same valence as the initial associative (priming) information, so that no attitude change would occur, and in other conditions the information was opposite in valence to the initial information, so that attitude change was induced. After these inductions, participants were exposed to a final set of information relevant to the target person that comprised either strong or weak arguments.

There were several notable findings from this study. First, Rydell et al. (2008) found that attitude change was associated with increased information processing as evidenced by a greater impact of argument quality on attitudes, replicating the Petty et al. (2006) results. Furthermore, attitude change was associated with greater reports of general discomfort regarding the attitude object compared to no change. Of most interest, Rydell and colleagues (2008) showed that the measure of general discomfort mediated the observed information-processing effect, supporting the idea that implicit ambivalence is uncomfortable and enhances information processing. This discomfort does not result from any explicit conflict, however. Rather, the discomfort stems from the conflict between an endorsed and an unendorsed evaluative association regarding the attitude object.

Implicit Ambivalence in Racial Attitudes

The research described so far provides convergent evidence for the viability of the implicit ambivalence idea. Another area of research that is consistent with this notion is in the domain of prejudice in racial attitudes. There are now a number of studies suggesting that European American (White) individuals sometimes engage in greater processing of a persuasive message when it comes from a Black person rather than a White person (e.g., White & Harkins, 1994). Petty, Fleming, and White (1999) suggested that this enhanced scrutiny might be due to the motive of Whites to guard against possible prejudice (from themselves or others) toward Black sources (i.e., watchdog motivation). Thus, they found that individuals low in prejudice were most likely to show the enhanced scrutiny effect, because they presumably were most concerned about scrutinizing information from a Black source to avoid showing any bias. The enhanced scrutiny also occurs when the information is about a Black target rather than from a Black source (Fleming, Petty, & White, 2005).

In a recent series of studies, we tested a variation of the watchdog hypothesis based on the idea of implicit ambivalence (Johnson, Petty, Briñol, See, & Fleming, 2011b). Specifically, the implicit ambivalence notion suggests that among individuals low in explicit (i.e., deliberative) prejudice, it

is those who are also high in implicit (i.e., automatic) prejudice who will do the most processing; that is, people who do not want to be prejudiced, or who see themselves as unprejudiced (low explicit prejudice) but harbor automatic negative reactions toward Blacks (high implicit prejudice), should experience conflict and be the most vigilant in guarding against prejudice (Devine, Monteith, Zuwerink, & Elliot, 1991; Monteith, 1993; Monteith et al., 1993). One way to guard against prejudice is to engage in careful processing of messages from Blacks. Although this prediction is consistent with the initial watchdog notion, the implicit ambivalence framework also makes a novel prediction that individuals who are high in explicit prejudice, but low in implicit prejudice, would also engage in enhanced information processing. In watchdog terminology, they would want to be sure that they demonstrated desired levels of prejudice. In the implicit ambivalence terminology, the enhanced processing stems from the discomfort associated with being ambivalent at an implicit rather than explicit level.

To examine the implicit ambivalence idea in the domain of prejudice, a recent study used both automatic and deliberative measures to assess Ohio State University students' attitudes toward African Americans (Johnson et al., 2011b). The automatic measure was an IAT in which stereotypically Black names (e.g., Tyrone, LaToya) and White names (e.g., Andrew, Katie) were paired with good (e.g., *freedom, love*) and bad (e.g., *poison, disease*) terms (see Greenwald et al., 1998). The explicit measure comprised a series of statements about Blacks (e.g., "On the whole, Black people do not stress education or training") to which participants rated their extent of agreement (see Katz & Hass, 1988). The explicit and implicit measures of attitudes were unrelated to each other. Following previous research on explicit–implicit divergence (Briñol et al., 2006), an index of explicit–implicit discrepancy was formed as the absolute value of the difference between the standardized explicit and implicit measures of racial attitudes. The direction of discrepancy (i.e., implicit score more prejudiced than explicit, or vice versa) was also examined to see whether it mattered.

After completing the implicit and explicit measures of racial attitudes, all of the students were exposed to a message advocating a new program to hire African American faculty at their university that was supported with either strong or weak arguments. As in past research, the strong arguments were designed to elicit favorable thoughts if people thought about them, whereas the weak arguments were designed to elicit mostly unfavorable thoughts (see Petty & Cacioppo, 1986). Consistent with the idea that people with automatic–deliberative discrepancies would act as if they were ambivalent, discrepancy interacted with argument quality to predict attitudes toward the program; that is, as the discrepancy between racial attitudes assessed with implicit and explicit measures increased, attitudes toward the program were more affected by argument quality. The direction of the discrepancy did not further qualify the results. These findings indicate that among participants who were low in explicit prejudice, it was primarily those high in

implicit prejudice who engaged in greater scrutiny of a message about a program favoring Blacks, consistent with the original watchdog idea (Petty et al., 1999). However, among participants high in explicit prejudice, those who were low in implicit prejudice engaged in the greatest scrutiny. Overall, then, the results are quite compatible with the predictions derived from the implicit ambivalence framework.

Notably, Johnson and colleagues (2011b) provided participants' with a topic quite relevant to race—a persuasive message about the hiring of African American faculty. According to the implicit ambivalence notion, any element of a persuasive proposal relevant to one's automatic–deliberative racial discrepancies could be sufficient to trigger the discomfort that results in increased processing. For example, if instead of having a message related to Blacks, participants were exposed to a message presented by a Black source, implicit ambivalence would make the same prediction. To explore this possibility, in another study, Ohio State students were given a message on the race-irrelevant topic of supporting the use of phosphate-based laundry detergents. The message was presented by either a Black or a White communicator (Johnson et al., 2011b). Source race was manipulated by the presentation of a photo of either a Black or White male that accompanied the persuasive message participants read. The message was designed to argue unambiguously in favor of phosphate detergents and presented either strong or weak arguments in support of the proposal (see Briñol, Petty, & Tormala, 2004). In addition to measuring participants' attitudes toward phosphate detergents after reading the persuasive message, implicit and explicit racial attitudes were also assessed. The implicit measure was an IAT in which photos of Black and White faces were paired with evaluative stimuli (e.g., the terms vomit, freedom; see Greenwald et al., 1998). For the explicit measure, participants provided their extent of agreement with various statements about Blacks (see Katz & Hass, 1988). In accord with the first study, the explicit and implicit measures of racial attitudes were unrelated to each other, and an index of explicit–implicit discrepancy was formed as the absolute value of the difference between the standardized explicit and implicit measures.

Consistent with the idea that people engage in greater information processing when a message is relevant to their implicit–explicit discrepancies, a significant three-way interaction of source race, argument quality, and discrepancy was obtained. Decomposing the interaction as a function of source race revealed that discrepancy only led to an enhanced impact of argument quality on attitudes toward phosphate detergents when the source of the persuasive message was Black. Furthermore, consistent with the first study, the direction of the discrepancy did not further qualify the results. These findings suggest that individuals who are implicitly ambivalent—high in implicit but low in explicit prejudice, or the reverse—engage in greater information processing connected to the discrepancy (i.e., message from a Black source). In summary, much like explicit ambivalence, ambivalence at the

implicit level also leads to greater information processing, presumably in the service of reducing the discomfort that results from the discrepancy.

Meaningful Negation in Racial Attitudes

In line with the MCM, the research described so far reveals that when people doubt the validity of a stored association (and therefore do not use it when deliberatively responding), that association can still create what we have called implicit ambivalence and thereby have an impact on information processing. Thus, although people might not report being explicitly ambivalent about a given object or group, the discomfort associated with the object of implicit ambivalence is nonetheless consequential.

Although we have reviewed research showing that people who are implicitly ambivalent engage in greater information processing than those who are not, we have not shown that this processing is effective in removing the ambivalence. Indeed, it might *not* be effective, because ambivalent individuals are likely to aim to bolster their endorsed evaluation, leaving the negated association in tact. One might wonder, then, whether implicit ambivalence can be undermined by strengthening the negation tag for the nonendorsed association, so that it would be activated automatically. Interestingly, recent research suggests that this may be difficult. Indeed, some available research indicates that attempts to negate one's prejudice can backfire and actually strengthen the unfavorable evaluation. In particular, in one highly relevant study, Gawronski, Deutsch, Mbirkou, Seibt, and Strack (2008) attempted to train participants either to negate information that was consistent with the Black stereotype (i.e., hitting the space bar on a keyboard to indicate "no" whenever a Black person was paired with a negative trait) or to affirm stereotype-inconsistent information (i.e., hitting the space bar to indicate "yes" whenever a Black person was paired with a positive trait). This research found that only affirming stereotype-inconsistent information was effective in undermining automatic prejudice. In contrast, attempting to negate prejudice actually made the link between Black and bad stronger, enhancing automatic prejudice.

However, the MCM suggests the possibility that continual practice in negating unfavorable traits can lead to a stored negation tag and could at some point lead the person immediately to think "no" when negative traits are activated, undermining prejudice (e.g., "not lazy" is automatically activated and seen as positive; see also Maddux, Barden, Brewer, & Petty, 2005). Is there any evidence that practice in negation can be successful? In a highly relevant series of studies, Deutsch et al. (2006) had participants practice negating relatively novel positive or negative words and timed how long it took for them to indicate the reversed meaning of the phrase (e.g., "no disease" is positive and "no party" is negative) compared to affirmations (e.g., "party" is positive or "disease" is negative). The key result was that initially people took significantly longer to indicate the correct evalu-

ative meaning of a negation than an affirmation. However, after numerous practice trials, people became as fast to process the meaning of negations as they were initially to recognize the meaning of affirmations. In a second set of studies, Deutsch et al. compared the meaning of negations on both explicit and implicit measures. On explicit measures, negations were easily understood, but on a measure of automatic evaluation, words negated relatively infrequently showed the same evaluation as the affirmation (e.g., "no disease" activated as much negativity as "disease"). However, for more frequently encountered negations (e.g., "no problem"), the intended meaning was apparent on an automatic measure. Although this is a correlational study, Deutsch et al. argued that this result is consistent with the idea that with very extensive practice, negation can be successful in reversing specific word meanings. Thus, the Deutsch et al. studies suggest that, with practice, it might be possible for negation training to undermine specific stereotypes; that is, continued practice in thinking that Blacks are *not* lazy could mean that if "lazy" were activated in association with Black, it could be negated sufficiently quickly that its valence would reverse, producing a reduction in automatic prejudice.

Given that there is some suggestive evidence that negation can be successful in reversing the meaning of negative terms with extensive practice, why has the existing evidence with respect to the negation of racial stereotypes been unsupportive? In a recent series of studies, Johnson, Kopp, and Petty (2011) hypothesized that perhaps the negation training used in prior research (e.g., Gawronski et al., 2008) did not work because either there was insufficient practice or the negation used (i.e., "No") was somewhat ambiguous or insufficiently meaningful or powerful. To examine this idea, Johnson et al. (2011a) adapted a design from similar research on negating prejudice (see Gawronski et al., 2008; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000) but used a more powerful negation. Specifically, participants were trained to hit the space bar on their computer keyboard and think "No! That's wrong!" whenever they saw either a Black face paired with a negative trait word (e.g., poor) or a White face paired with a positive word (e.g., rich). For combinations inconsistent with the cultural stereotype of Blacks and Whites, participants were instructed to do nothing—simply to wait for the next face–trait combination to appear. In a second condition, the negate stereotype-inconsistent condition, participants' were trained to do just the opposite. Specifically, participants' were told to hit the spacebar and think "No! That's wrong!" each time they saw a Black face paired with a positive trait or a White face paired with a negative trait. Furthermore, for trials in which Black faces were paired with negative traits or White faces were paired with positive traits, participants were instructed to give no response and wait for the next face–trait combination to appear. Each participant practiced the training extensively (i.e., five blocks of 40 trials, resulting in 200 total trials). Just prior to completing negation training, participants responded to the same subliminal evaluative priming task used to assess automatic racial

attitudes by Gawronski et al. (2008; see Fazio et al., 1995). They responded to this same measure once again, immediately after training.

The results, depicted in Figure 9.2, indicate that participants who engaged in negation training of stereotype-consistent information showed a decrease in automatic racial prejudice following training. Conversely, for those participants who engaged in negation of stereotype-inconsistent information, an increase in automatic racial prejudice followed training. This pattern resulted in a significant two-way interaction of condition and time of measurement. These results suggest that negation training can be used as a tool to undermine the validity with which certain associations are held and can, in turn, modify automatic attitudes. Furthermore, given the differential change in automatic racial prejudice observed for the stereotype consistent and inconsistent training conditions from time 1 to time 2, it seems unlikely that negation is solely an online process. In order for the negation training to impact participants' automatic racial prejudice, it seems likely that the negation was encoded during the training and later retrieved spontaneously. Thus, consistent with the MCM, the present work supports the idea that negations can be stored along with evaluative associations in memory for later retrieval (Mayo, Schul, & Burnstein, 2004). Future research should examine the consequences of the storage of validity tags for implicit ambivalence and information processing; that is, when people store the rejection of their associations and this negation is activated automatically, the negated associations might not create implicit ambivalence anymore, thereby reducing the motivation to process discrepancy relevant information.

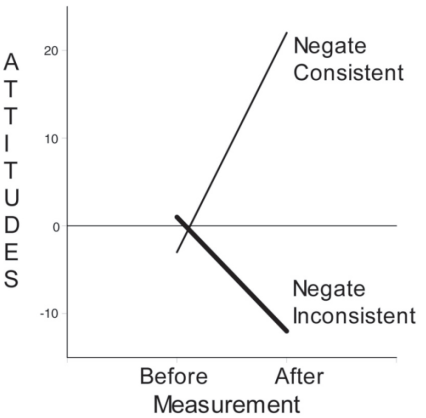

FIGURE 9.2. Automatic racial attitudes before and after negation training to prejudice-consistent and -inconsistent information. Data from Johnson et al. (2011).

CONCLUSION

In this chapter, we have argued that sometimes people can be *implicitly ambivalent*, even when they do not acknowledge that there are two valid sides to an issue (i.e., there is no explicit ambivalence). This prediction is based on the MCM of attitude structure, according to which positive and negative associations to an attitude object can be stored separately, along with metacognitive labels of validity. The research covered in this chapter has revealed that when people deny one evaluative reaction as invalid or believe it stems from some source other than their personal views, they do not use it in their deliberative self-reports, though it can still affect automatic evaluations. Furthermore, our research has shown that despite no impact on measures of explicit ambivalence, discrepancies between implicit and explicit attitudes and attitude change can result in an implicit ambivalence that is consequential. Unlike implicit balance that can lead to spontaneous changes in attitudes, implicit ambivalence is associated with changes that result from enhanced processing of relevant information. We concluded with research suggesting that practice in negation might eventually reduce implicit ambivalence, though additional research is needed to confirm this possibility.

NOTES

1. It is also possible to extend balance theory to more than three cognitive elements (e.g., see Insko, 1984, for more details on balance principles).
2. People can deal with ambivalence in ways other than enhanced thinking. These include changing discrepant elements (e.g., Festinger, 1957; Harmon-Jones & Mills, 1999), minimizing the salience or importance of the inconsistency (e.g., Steele & Liu, 1981), and ignoring or defensively avoiding discrepancy-related information. Although different mechanisms of reducing inconsistency might be substitutable for each other (e.g., Tesser, 2001), thinking about target (and ambivalence diagnostic) information is a good strategy for discrepancy reduction, because it can provide individuals with a solid and stable basis for their attitudes.

REFERENCES

Aaker, J. (1999). The malleable self: The role of self-expression in persuasion. *Journal of Marketing Research, 36*, 45–57.

Abelson, R. P., Aronson, E., McGuire, W. J., Newcomb, M., Rosenberg, M. J., & Tannenbaum, P. H. (1968). *Theories of cognitive consistency: A sourcebook*. Chicago: Rand McNally.

Abelson, R. P., & Rosenberg, M. J. (1958). Symbolic psycho-logic: A model of attitudinal cognition. *Behavioral Science, 3*, 1–13.

Armitage, C. J., & Conner, M. (2000). Attitudinal ambivalence: A test of three key hypotheses. *Personality and Social Psychology Bulletin, 26*, 1421–1432.

Aronson, E. (1969). Some antecedents of interpersonal attraction. *Nebraska Symposium on Motivation, 17*, 143–173.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of automatic attitude activation effect. *Journal of Personality and Social Psychology, 6*, 893–912.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology, 32*, 104–128.

Belk, R. W. (1988). Possessions and the extended self. *Journal of Consumer Research, 15*, 139–168.

Briñol, P., & DeMarree, K. G. (Eds.). (in press). *Social metacognition*. New York: Psychology Press.

Briñol, P., Petty, R., & Tormala, Z. (2004). Self-validation of cognitive responses to advertisements. *Journal of Consumer Research, 30*, 559–573.

Briñol, P., Petty, R. E., & Wheeler, S. C. (2006). Discrepancies between explicit and implicit self-concepts: Consequences for information processing. *Journal of Personality and Social Psychology, 91*, 154–170.

Brunel, F. F., Tietje, B. C., & Greenwald, A. G. (2004). Is the Implicit Association Test a valid and valuable measure of implicit consumer social cognition? *Journal of Consumer Psychology, 14*, 385–404.

Byrne, D. (1971). *The attraction paradigm*. New York: Academic Press.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review, 1*, 3–25.

Clark, J. K., Wegener, D. T., & Fabrigar, L. (2008). Attitudinal ambivalence and message-based persuasion: Motivated processing of proattitudinal information and avoidance of counterattitudinal information. *Personality and Social Psychology Bulletin, 34*, 565–577.

Clarkson, J. J., Tormala, Z. L., & Rucker, D. D. (2008). A new look at the consequences of attitude certainty: The amplification hypothesis. *Journal of Personality and Social Psychology, 95*, 810–825.

Converse, P. E. (1970). Attitudes and non-attitudes: Continuation of a dialogue. In E. R. Tufte (Ed.), *The quantitative analysis of social problems* (pp. 168–189). Reading, MA: Addison-Wesley.

DeCoster, J., Banner, M. J., Smith, E. R., & Semin, G. R. (2006). On the inexplicability of the implicit: Differences in the information provided by implicit and explicit tests. *Social Cognition, 24*, 5–21.

de Liver, Y., van der Pligt, J., & Wigboldus, D. (2007). Positive and negative associations underlying ambivalent attitudes. *Journal of Experimental Social Psychology, 43*, 319–326.

Dempsey, M. A., & Mitchell, A. A. (2010). The influence of implicit attitudes on choice when consumers are confronted with conflicting attribute information. *Journal of Consumer Research, 37*, 614–625.

Deutsch, R., Gawronski, B., & Strack, F. (2006). At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology, 91*, 385–405.

Devine, P., Monteith, M., Zuwerink, J., & Elliot, A. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology, 60*, 817–830.

Dovidio, J., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the

nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33*, 510–540.

Epstein, S. (2003). Cognitive-experiential self-theory of personality. In T. Millon & M. Lerner (Eds.), *Handbook of psychology: Personality and social psychology* (Vol. 5, pp. 159–184). Hoboken, NJ: Wiley.

Fazio, R. H. (1995). Attitudes as object–evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–283). Hillsdale, NJ: Erlbaum.

Fazio, R. H. (2007). Attitudes as object–evaluation associations of varying strength. *Social Cognition, 25*, 603–637.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013–1027.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229–238.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fiske, S. T., & Pavelchak, M. A. (1986). Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 167–203). New York: Guilford Press.

Fleming, M. A., Petty, R. E., & White, P. (2005). Stigmatized targets and evaluation: Prejudice as a determinant of attribute scrutiny and polarization. *Personality and Social Psychology Bulletin, 4*, 496–507.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692–731.

Gawronski, B., Bodenhausen, G. V., & Becker, A. (2007). I like it, because I like myself: Associative self-anchoring and post-decisional change of implicit evaluations. *Journal of Experimental Social Psychology, 43*, 221–232.

Gawronski, B., Deutsch, R., Mbirkou, S., Seibt, B., & Strack, F. (2008). When "Just say no" is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation. *Journal of Experimental Social Psychology, 44*, 370–377.

Gawronski, B., & Payne, B. K. (Eds.). (2010). *Handbook of implicit social cognition: Measurement, theory, and applications*. New York: Guilford Press.

Gawronski, B., Walther, E., & Blank, H. (2005). Cognitive consistency and the formation of interpersonal attitudes: Cognitive balance affects the encoding of social information. *Journal of Experimental Social Psychology, 41*, 618–626.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46*, 107–119.

Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology, 59*, 601–613.

Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology, 65*, 221–233.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mel-

lott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3–25.

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79*, 1022–1038.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Task. *Journal of Personality and Social Psychology, 74*, 1464–1480.

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology, 90*, 1–20.

Gross, S. R., Holtz, R., & Miller, N. (1995). Attitude certainty. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 215–245). Hillsdale, NJ: Erlbaum.

Hänze, M. (2001). Ambivalence, conflict, and decision making: Attitudes and feelings in Germany towards NATO's military intervention in the Kosovo War. *European Journal of Social Psychology, 31*, 693–706.

Harmon-Jones, E., & Mills, J. (1999). An introduction to cognitive dissonance theory and an overview of current perspectives on the theory. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive dissonance: Progress on a pivotal theory in social psychology* (pp. 3–21). Washington, DC: American Psychological Association.

Hass, R. G., Katz, I., Rizzo, N., Bailey, J., & Moore, L. (1992). When racial ambivalence evokes negative affect, using a disguised measure of mood. *Personality and Social Psychology Bulletin, 18*, 786–797.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*, 319–340.

Hodson, G., Maio, G. R., & Esses, V. M. (2001). The role of attitudinal ambivalence in susceptibility to consensus information. *Basic and Applied Social Psychology, 23*, 197–205.

Horcajo, J., Briñol, P., & Petty, R. E. (2010). Consumer persuasion: Indirect change and implicit balance. *Psychology and Marketing, 27*, 938–963.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. New Haven, CT: Yale University Press.

Insko, C. A. (1984). Balance theory, the Jordan paradigm, and the Wiest tetrahedron. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 89–140). San Diego: Academic Press.

James, W. (1890). *Principles of psychology*. Chicago: Encyclopedia Britannica.

Johnson, I., Kopp, B., & Petty, R. E. (2011a). *Just say "No" (and mean it): Meaningful negation and the modification of automatic racial attitudes*. Unpublished manuscript, Ohio State University, Columbus, OH.

Johnson, I. R., Petty, R. E., Briñol, P., See, Y. H. M., & Fleming, M. (2011b). *Racial scrutiny as a function of automatic-deliberative discrepancies in racial attitudes*. Unpublished manuscript, Ohio State University, Columbus, OH.

Jonas, K., Diehl, M., & Bromer, P. (1997). Effects of attitudinal ambivalence on information processing and attitude–intention consistency. *Journal of Experimental Social Psychology, 33*, 190–210.

Jost, J. T., Kruglanski, A. W., & Nelson, T. O. (1998). Social metacognition: An expansionist review. *Personality and Social Psychology Review, 2*, 137–154.

Kaplan, K. J. (1972). On the ambivalence–indifference problem in attitude theory and measurement: A suggested modification of the semantic differential technique. *Psychological Bulletin, 77*, 361–372.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology, 55*, 893–905.

Katz, I., Wackenhut, J., & Hass, R. G. (1986). Racial ambivalence, value duality, and behavior. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 35–59). San Diego: Academic Press.

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78*, 871–888.

Kleine, R. E., III, Kleine, S. S., & Kernan, J. B. (1993). Mundane consumption and the self: A social identity perspective. *Journal of Consumer Psychology, 2*, 209–235.

Langer, T., Walther, E., Gawronski, B., & Blank, H. (2009). When linking is stronger than thinking: Associative transfer of valence disrupts the emergence of cognitive balance after attitude change. *Journal of Experimental Social Psychology, 45*, 1232–1237.

Maddux, W. W., Barden, J., Brewer, M. B., & Petty, R. E. (2005). Saying no to negativity: The effects of context and motivation to control prejudice on automatic evaluative responses. *Journal of Experimental Social Psychology, 41*, 19–35.

Maio, G. R., Bell, D. W., & Esses, V. M. (1996). Ambivalence in persuasion: The processing of messages about immigrant groups. *Journal of Experimental Social Psychology, 32*, 513–536.

Mayo, R., Schul, Y., & Burnstein, E. (2004). I am not guilty vs. I am innocent: Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology, 40*, 433–449.

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. *Journal of Personality and Social Psychology, 65*, 469–485.

Monteith, M., Devine, P., & Zuwerink, J. (1993). Self-directed versus other-directed affect as a consequence of prejudice-related discrepancies. *Journal of Personality and Social Psychology, 64*, 198–210.

Newby-Clark, I. R., McGregor, I., & Zanna, M. P. (2002). Thinking and caring about cognitive inconsistency: When and for whom does attitudinal ambivalence feel uncomfortable? *Journal of Personality and Social Psychology, 82*, 152–166.

Newcomb, T. M. (1978). The acquaintance process: Looking mainly backward. *Journal of Personality and Social Psychology, 36*, 1075–1083.

Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review 62*, 42–55.

Peters, K. R., & Gawronski, B. (2011). Are we puppets on a string? Comparing the impact of contingency and validity on implicit and explicit evaluations. *Personality and Social Psychology Bulletin, 37*, 557–569.

Petty, R. E. (2006). A metacognitive model of attitudes. *Journal of Consumer Research, 33*, 22–24.

Petty, R. E., & Briñol, P. (2006). A metacognitive approach to "implicit" and "explicit" evaluations: Comment on Gawronski and Bodenhausen (2006). *Psychological Bulletin, 132*, 740–744.

Petty, R. E., & Briñol, P. (2009). Implicit ambivalence: A meta-cognitive approach. In

R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 119–164). New York: Psychology Press.

Petty, R. E., Briñol, P., & DeMarree, K. G. (2007). The meta-cognitive model (MCM) of attitudes: Implications for attitude measurement, change and strength. *Social Cognition, 25*, 657–686.

Petty, R. E., Briñol, P., Tormala, Z. L., & Wegener, D. T. (2007). The role of metacognition in social judgment. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of principles* (2nd ed., pp. 254–284). New York: Guilford Press.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.

Petty, R. E., & Fazio, R. H., & Briñol, P. (2009a). The new implicit measures: An overview. In *Attitudes: Insights from the new implicit measures* (pp. 3–18). New York: Psychology Press.

Petty, R. E., Fazio, R. H., & Briñol, P. (Eds.). (2009b). *Attitudes: Insights from the new implicit measures*. New York: Psychology Press.

Petty, R. E., Fleming, M. A., & White, P. H. (1999). Stigmatized sources and persuasion: Prejudice as a determinant of argument scrutiny. *Journal of Personality and Social Psychology, 76*, 19–34.

Petty, R. E., Tormala, Z. L., Briñol, P., & Jarvis, W. B. G. (2006). Implicit ambivalence from attitude change: An exploration of the PAST Model. *Journal of Personality and Social Psychology, 90*, 21–41.

Priester, J. R., & Petty, R. E. (1996). The gradual threshold model of ambivalence: Relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of Personality and Social Psychology, 71*, 431–449.

Priester, J. R., Wegener, D., Petty, R. E., & Fabrigar, L. (1999). Examining the psychological process underlying the sleeper effect: The elaboration likelihood model explanation. *Media Psychology, 1*, 27–48.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32*, 880–892.

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology, 91*, 995–1008.

Rydell, R. J., McConnell, A. R., & Mackie, D. M. (2008). Consequences of discrepant explicit and implicit attitudes: Cognitive dissonance and increased information processing. *Journal of Experimental Social Psychology, 44*, 1526–1532.

Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of two minds: Forming and changing valence-inconsistent implicit and explicit attitudes. *Psychological Science, 17*, 954–958.

Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition, 25*, 638–656.

Sparks, P., Harris, P. R., & Lockwood, N. (2004). Predictors and predictive effects of ambivalence. *British Journal of Social Psychology, 43*, 371–383.

Staats, A. W., & Staats, C. K. (1958). Attitudes established by classical conditioning. *Journal of Abnormal and Social Psychology, 11*, 187–192.

Steele, C., & Liu, T. (1981). Making the dissonant act unreflective of self: Dissonance

avoidance and the expectancy of a value-affirming response. *Personality and Social Psychology Bulletin, 7*, 393–397.

Tesser, A. (2001). On the plasticity of self-defense. *Current Directions in Psychological Science, 10*, 66–69.

Thompson, M. M., Zanna, M. P., & Griffin, D. W. (1995). Let's not be indifferent about (attitudinal) ambivalence. In *Attitude strength: Antecedents and consequences*. Mahwah, NJ: Erlbaum.

Tourangeau, R., Rasinski, K. A., Bradburn, N., & D'Andrade, R. (1989). Belief accessibility and context effects in attitude measurement. *Journal of Experimental Social Psychology, 25*, 401–421.

van Harreveld, F., van der Pligt, J., & de Liver, Y. N. (2009). The agony of ambivalence and ways to resolve it. *Personality and Social Psychology Review, 13*, 45–61.

Walther, E. (2002). Guilty by mere association: Evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology, 82*, 919–934.

Walther, E., & Trasselli, C. (2003). I like her, because I like myself: Self-evaluation as a source of interpersonal attitudes. *Experimental Psychology, 50*, 239–246.

White, P. H., & Harkins, S. G. (1994). Race of source effects in the elaboration likelihood model. *Journal of Personality and Social Psychology, 67*, 790–807.

Wicklund, R. A., & Gollwitzer, P. M. (1982). *Symbolic self-completion*. Hillsdale, NJ: Erlbaum.

Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In L.L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 37–65). Hillsdale, NJ: Erlbaum.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107*, 101–126.

Wittenbrink, B., & Schwarz, N. (Eds.). (2007). *Implicit measures of attitudes*. New York: Guilford Press.

Wyer, N. (2010). You never get a second chance to make a first (implicit) impression: The role of elaboration in the formation and revision of implicit impressions. *Social Cognition, 28*, 1–19.